

OIPF

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,739

DATE: 04/12/2001

TIME: 16:35:45

Input Set : N:\Crif3\RULE60\09784739.txt

Output Set : N:\CRF3\04192001\I784739.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Goli, Surya K.
8 Hillman, Jennifer L.

C--> 10 (ii) TITLE OF INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE

12 (iii) NUMBER OF SEQUENCES: 5

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
16 (B) STREET: 3174 Porter Drive
17 (C) CITY: Palo Alto
18 (D) STATE: CA
19 (E) COUNTRY: US
20 (F) ZIP: 94304

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette
24 (B) COMPUTER: IBM Compatible
25 (C) OPERATING SYSTEM: DOS
26 (D) SOFTWARE: FastSEQ Version 2.0

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/784,739

C--> 30 (B) FILING DATE: 14-Feb-2001

32 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/309,320

35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J.
39 (B) REGISTRATION NUMBER: 36,749
40 (C) REFERENCE/DOCKET NUMBER: PF-0162 US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 415-855-0555

44 (B) TELEFAX: 415-845-4166

45 (C) TELFX:

48 (2) INFORMATION FOR SEQ ID NO: 1:

50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 222 amino acids

52 (B) TYPE: amino acid

53 (C) ORGANISM: Homo

54 (D) TAXID: 9606

55 (E) SOURCE:

56 (A) LIBRARY: Consensus

57 (B) CLONE: Consensus

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

62 Met Ala Ala Arg Pro Lys Leu His Tyr Pro Asn Gly Arg Gly Arg Met

63 1 15

64 Val Val Val Arg Ile Val Leu Ala Ala Ala Gly Val His Pro Asn Met

65 16 31

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,739

DATE: 04/19/2001

TIME: 16:35:45

Input Set : N:\Crif3\RULE60\09784739.txt

Output Set : N:\CRF3\04192001\1784739.raw

```

66  Glu Phe Leu Glu Thr Lys Gln Gln Leu Tyr Lys Leu Gln Asp Gly Asn
67      35              40              45
68  His Leu Leu Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys
69      50              55              60
70  Leu Val Gln Thr Arg Ser Ile Leu His Tyr Ile Ala Asp Lys His Asn
71      65              70              75              80
72  Leu Phe Gly Lys Asn Leu Lys Glu Arg Thr Leu Ile Asp Met Tyr Val
73      85              90              95
74  Glu Gly Thr Leu Asp Leu Leu Glu Leu Leu Ile Met His Pro Phe Leu
75      100             105             110
76  Lys Pro Asp Asp Gln Gln Lys Glu Val Val Asn Met Ala Gln Lys Ala
77      115             120             125
78  Ile Ile Arg Tyr Phe Pro Val Phe Glu Lys Ile Leu Arg Gly His Gly
79      130             135             140
80  Gln Ser Phe Leu Val Gly Asn Gln Leu Ser Leu Ala Asp Val Ile Leu
81      145             150             155             160
82  Leu Gln Thr Ile Leu Ala Leu Glu Glu Lys Ile Pro Asn Ile Leu Ser
83      165             170             175
84  Ala Phe Pro Phe Leu Gln Glu Tyr Thr Val Lys Leu Ser Asn Ile Pro
85      180             185             190
86  Thr Ile Lys Arg Phe Leu Glu Pro Gly Ser Lys Lys Lys Pro Pro Pro
87      195             200             205
88  Asp Glu Ile Tyr Val Arg Thr Val Tyr Asn Ile Phe Arg Pro
89      210             215             220
90 (2) INFORMATION FOR SEQ ID NO: 2:
91 (i) SEQUENCE CHARACTERISTICS:
92 (A) LENGTH: 300 base pairs
93 (B) TYPE: nucleic acid
94 (C) STRANDEDNESS: single
95 (D) TOPOLOGY: linear
96 (vii) IMMEDIATE SOURCE:
97 (A) LIBRARY: Consensus
98 (B) CLONE: Consensus
99 (x.) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
105 ATGCCAGCAA GGCCCAAGCT CCACTATCCC AACGGAAGAG GCCGGATGGA GTCCGTGAGA 60
106 TGGGTTTTAG CTGCCGCCGG AGTCGAGTTT GATGAAGAAT TTCTGAAAC AAAAGAACAG 120
107 TTGTACAAGT TGCACGATGG TAACCACCTG CTGTCCAAC AAGTGCCCAT GGTGAAATT 180
108 GACGGGATGA AGTTGGTACA GACCCGAAGC ATTCTCCACT ACATAGCAGA CAAGCACAAT 240
109 CTTCTTGGCA AGAAGCTCAA GGAGAGAACC CTGATTGACA TGTACGTGGA TGGGACACTG 300
110 AATTTATGG AACTTTTAT CATGATATT TTTTAAAAA CAGATGATTA TAAAAAAGAA 360
111 TATTTAAAT TTTTAAATA ATATAAAT ATATATTTT CTGTTTAA AAAATTTTA 420
112 AATTTTAA ATATAAATTT TTTTATTTT AATCAGTTCA GTTTTAAAT TTTTATTTA 480
113 TTTTAAATA TTTTATTTT AATAAATAA ATTTTAAAT TTTTATTTT AATTTTATTT 540
114 TTTTAAAT ATATAAAT AATAAATAA ATTTTAAAT TTTTATTTT AATTTTATTT 600
115 TTTTAAATA AATAAATAA ATTTTAAAT TTTTATTTT AATTTTATTT AATTTTATTT 660
116 AATAAATAA AATAAATAA ATTTTAAAT TTTTATTTT AATTTTATTT AATTTTATTT 720
117 TATATATATG TCTTAAATGA TCCAGCTCT GTCATGGTGC TATCTATGTA TAAATTTGGG 780
118 TCTAAGTTG GGTCTTTTGT
119 (2) INFORMATION FOR SEQ ID NO: 3:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,739

DATE: 04/19/2001

TIME: 16:35:45

Input Set : N:\Crf3\RULE60\09784739.txt

Output Set : N:\CRF3\04192001\I784739.raw

```

121      (i) SEQUENCE CHARACTERISTICS:
122          (A) LENGTH: 222 amino acids
123          (B) TYPE: amino acid
124          (C) STRANDEDNESS: single
125          (D) TOPOLOGY: linear
126      (ii) IMMEDIATE SOURCE:
127          (A) LIBRARY: GenBank
128          (B) CLONE: 825605
129      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
130 Met Ala Glu Lys Pro Lys Leu His Tyr Ser Asn Thr Arg Gly Arg Met
131 1 5 10 15
132 Gln Ser Ile Arg Trp Leu Leu Ala Ala Ala Gly Val Gln Phe Glu Glu
133 20 25 30
134 Lys Phe Ile Lys Ser Ala Glu Asp Leu Asp Lys Leu Arg Asn Asp Gly
135 35 40 45
136 Tyr Leu Met Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys
137 50 55 60
138 Leu Val Gln Thr Arg Ala Ile Leu Asn Tyr Ile Ala Ser Lys Tyr Asn
139 65 70 75 80
140 Leu Tyr Gly Lys Asp Ile Lys Glu Lys Ala Leu Ile Asp Met Tyr Ile
141 85 90 95
142 Glu Gly Ile Ala Asp Leu Gly Glu Met Ile Leu Leu Leu Pro Phe Thr
143 100 105 110
144 Gln Pro Glu Glu Gln Asp Ala Lys Leu Ala Leu Ile Gln Glu Lys Thr
145 115 120 125
146 Lys Asn Arg Tyr Phe Pro Ala Phe Glu Lys Val Leu Lys Ser His Gly
147 130 135 140
148 Gln Asp Tyr Leu Val Gly Asn Lys Leu Ser Arg Ala Asp Ile His Leu
149 145 150 155 160
150 Val Glu Leu Leu Tyr Tyr Val Glu Glu Leu Asp Ser Ser Leu Ile Ser
151 165 170 175
152 Ser Phe Pro Leu Leu Lys Ala Leu Lys Thr Arg Ile Ser Asn Leu Pro
153 180 185 190
154 Thr Val Lys Lys Phe Leu Gln Pro Gly Ser Pro Arg Lys Pro Pro Met
155 195 200 205
156 Asp Glu Lys Ser Leu Glu Glu Ser Arg Lys Ile Phe Arg Phe
157 210 215 220

```

163 (2) INFORMATION FOR SEQ ID NO: 4:

```

164      (i) SEQUENCE CHARACTERISTICS:
165          (A) LENGTH: 222 amino acids
166          (B) TYPE: amino acid
167          (C) STRANDEDNESS: single
168          (D) TOPOLOGY: linear
169      (ii) IMMEDIATE SOURCE:
170          (A) LIBRARY: GenBank
171          (B) CLONE: 259141
172      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
173 Met Ala Glu Lys Pro Lys Leu His Tyr Phe Asn Ala Ala Gly Arg Met
174 1 5 10 15

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,739

DATE: 04/19/2001

TIME: 16:35:45

Input Set : N:\Crif3\RULE60\09784739.txt

Output Set : N:\CRF3\04192001\I784739.raw

```

179  Ala Ser Thr Arg Trp Leu Leu Ala Ala Ala Gly Val Glu Ile Glu His
180                20                25                30
181  Lys Phe Ile Lys Ser Ala Glu Asp Leu Asp Lys Leu Arg Asn Asp Gly
182                35                40                45
183  Tyr Leu Met Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys
184                50                55                60
185  Leu Val Gln Thr Arg Ala Ile Leu Asn Tyr Ile Ala Ser Lys Tyr Asn
186                65                70                75                80
187  Leu Tyr Gly Lys Asp Ile Lys Glu Arg Ala Leu Ile Asp Met Tyr Ile
188                85                90                95
189  Glu Gly Ile Ala Asp Leu Gly Glu Met Ile Leu Leu Leu Pro Val Cys
190                100               105               110
191  Pro Pro Glu Glu Lys Asp Ala Lys Leu Ala Leu Ile Lys Glu Lys Ile
192                115               120               125
193  Lys Asn Arg Tyr Phe Pro Ala Phe Glu Lys Val Leu Lys Ser His Gly
194                130               135               140
195  Gln Asp Tyr Leu Val Gly Asn Lys Leu Ser Arg Ala Asp Ile His Leu
196                145               150               155               160
197  Val Glu Leu Leu Tyr Tyr Val Glu Glu Leu Asp Ser Ser Leu Ile Ser
198                165               170               175
199  Ser Phe Pro Leu Lys Ala Leu Lys Thr Arg Ile Ser Asn Leu Pro
200                180               185               190
201  Thr Val Lys Lys Phe Leu Gln Pro Gly Ser Pro Arg Lys Pro Pro Met
202                195               200               205
203  Asp Glu Lys Ser Leu Glu Glu Ala Arg Lys Ile Phe Arg Phe
204                210               215               220

```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 193710

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

219 Met Ala Ala Lys Pro Lys Leu Tyr Tyr Phe Asn Gly Arg Gly Arg Met
220      1      5      10      15
221 Glu Ser Ile Arg Trp Leu Leu Ala Ala Ala Gly Val Glu Phe Glu Glu
222      20      25      30
223 Asn Ile Ser Ala Thr Arg Val Thr Tyr Glu Lys Met Glu Lys Asp Ile
224      35      40      45
225 His Leu Leu Ile His Gly Glu Val Trp Leu Val Leu Ile Asp Gly Met Met
226      50      55      60
227 Leu Thr Glu Thr Arg Ala Ile Leu Ser Tyr Leu Ala Ala Lys Tyr Arg
228      65      70      75      80
229 Leu Tyr Gly Lys Asp Leu Lys Glu Arg Val Arg Ile Asp Met Tyr Ala
230      85      90      95
231 Asp Gly Thr Glu Asp Leu Met Met Met Ile Ala Val Ala Trp Phe Lys
232

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,739

DATE: 04/10/2001

TIME: 16:35:45

Input Set : N:\Crif3\RULE60\09784739.txt

Output Set : N:\CRF3\04192001\1784739.raw

```

233           100           105           110
234 Thr Pro Lys Glu Lys Glu Glu Ser Tyr Asp Leu Ile Leu Ser Arg Ala
235           115           120           125
236 Lys Thr Arg Tyr Phe Pro Val Phe Glu Lys Ile Leu Lys Asp His Gly
237           130           135           140
238 Glu Ala Phe Leu Val Gly Asn Gln Leu Ser Trp Ala Asp Ile Gln Leu
239 145           150           155           160
240 Leu Glu Ala Ile Leu Met Val Glu Glu Leu Ser Ala Pro Val Leu Ser
241           165           170           175
242 Asp Phe Pro Leu Leu Gln Ala Phe Lys Thr Arg Ile Ser Asn Ile Pro
243           180           185           190
244 Thr Ile Lys Lys Phe Leu Gln Pro Gly Ser Gln Arg Lys Pro Pro Pro
245           195           200           205
246 Asp Gly Pro Tyr Val Glu Val Val Arg Ile Val Leu Lys Phe
247           210           215           220

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/784,739

DATE: 04/18/2001

TIME: 16:35:46

Input Set : N:\CrF3\RULE60\09784739.txt

Output Set : N:\CRF3\04192001\1784739.raw

Line 5 Mail : Keyword misspelled or invalid format. [(1) GENERAL INFORMATION:]
Line 10 Mail : Keyword misspelled or invalid format. [(11) TITLE OF INVENTION:]
Line 29 Mail : Keyword misspelled or invalid format. [(A) APPLICATION NUMBER:]
Line 30 Mail : Keyword misspelled or invalid format. [(B) FILING DATE:]